Point-to-point responses

We appreciate the careful read-through and constructive comments from both reviewers. The reviewers' comments are copied in **black** with our point-to-point responses in **blue** and revised text in **red** or pasted.

Response to Reviewer #2:

The study presents a technically detailed implementation of a Lagrangian model adapted for use with column data, X-STILT, focusing on emission ratios of CO:CO2 on sub-city scales using XCO measurements from TROPOMI and XCO2 measurements from OCO-2 and OCO-3. The level of technical detail is quite good (with a few exceptions, see comments below), and it is nice to see that the adapted X-STILT code has been made publicly available. The approach is interesting, particularly in its attempt to not make use of emissions inventories as part of the prior information. However some potential sources of error have been overlooked, particularly with respect to the high- frequency variability of the CO:CO2 ratio, which seems to be assumed to be constant during the time between measurements. This is not something that the current study can really correct, but it needs to be included in the discussion as a clear limitation of the current results. Despite these limitations, the study is certainly appropriate for publication in ACP once these concerns are addressed. A good proof-reading of the paper is also needed before resubmission: while it was almost always clear what was meant, and the paper was well-written and easy to follow, there were lots of missing articles etc. that the native speakers among the co-authors could clear up quickly.

> We truly thank Julia Marshall for the detailed and constructive comments and have tried to address the main concerns about the high-frequency variability in ER_{CO}.

Major concerns:

The authors have gone to great pains to try to correct for temporal shifts between the measurements, considering the impact of the different meteorological conditions and averaging kernels. *What was not taken into account is changes in the CO:CO_2 ratio over the course of a day.* While this may not be as critical for heavy industry and power generation, other sectors (e.g. traffic) have highly heterogenous emission ratios in time, depending also on traffic patterns. Having the XCO and XCO_2 measurements offset by even a couple of hours complicates this approach considerably, and might also cloud the proposed analysis of temporal trends in the emission ratio over the year. That does not mean that nothing can be learnt from this approach, only that this neglected error source needs to be explicitly described. In any case, GeoCarb data will make such analyses considerably easier in the future.

> We agree with Julia that high-frequency variability in CO or CO_2 emissions may influence the ER_{CO}, although the overpass time difference has been limited to 1-3 hours. One can likely rely on bottom-up emission inventories or prior assumptions to infer hourly emission patterns. However, because emissions or combustion characteristics are what we are solving for, it is quite challenging to properly account for such mismatch WITHOUT involving prior assumptions towards emissions themselves. Such hourly emission variations depend on the relative contributions from individual emission sectors for a given city.
>
> We have now carried out a sensitivity analysis to investigate how such a mismatch in emission timing may affect the ER_{CO}, or more generally the high-frequency variability of ERCO (as reviewer #2 also mentioned). Thanks to the TCCON network that provides high-frequency XCO and XCO_2 measurements (TCCON 2022). We utilized the latest GGG2020 version with several upgrades including a much-improved prior profile. Because XCO and XCO_2 are simultaneously retrieved, we can

assume that atmospheric transport associated with two species are the same and their enhancement ratio fully reflected the emission characteristics once species-specific averaging kernel profiles are corrected for following Appendix A2 in Hedelius et al. (2018).

Here we reported the estimated enhancement ratio at Caltech with background observations from another TCCON site outside the LA basin (see figure and figure caption shown below). Because the temporal frequency of the Caltech TCCON site may not perfectly match that of the background TCCON site, we first averaged observations from both sites to each 20-min interval and calculated the enhancements and enhancement ratio. Indeed - the variation in ER_{CO} can be large throughout the day. Note that such variations may reflect not only the change in FF emissions (e.g., due to traffic) but also measurement noise (given these high-frequency data) as well as changes in the meteorology throughout the day (ocean versus mountain flow for LA). Nevertheless, changes in ER_{CO} between OCO-3 and TROPOMI times inferred from concurrent TCCON measurements appear to be small.

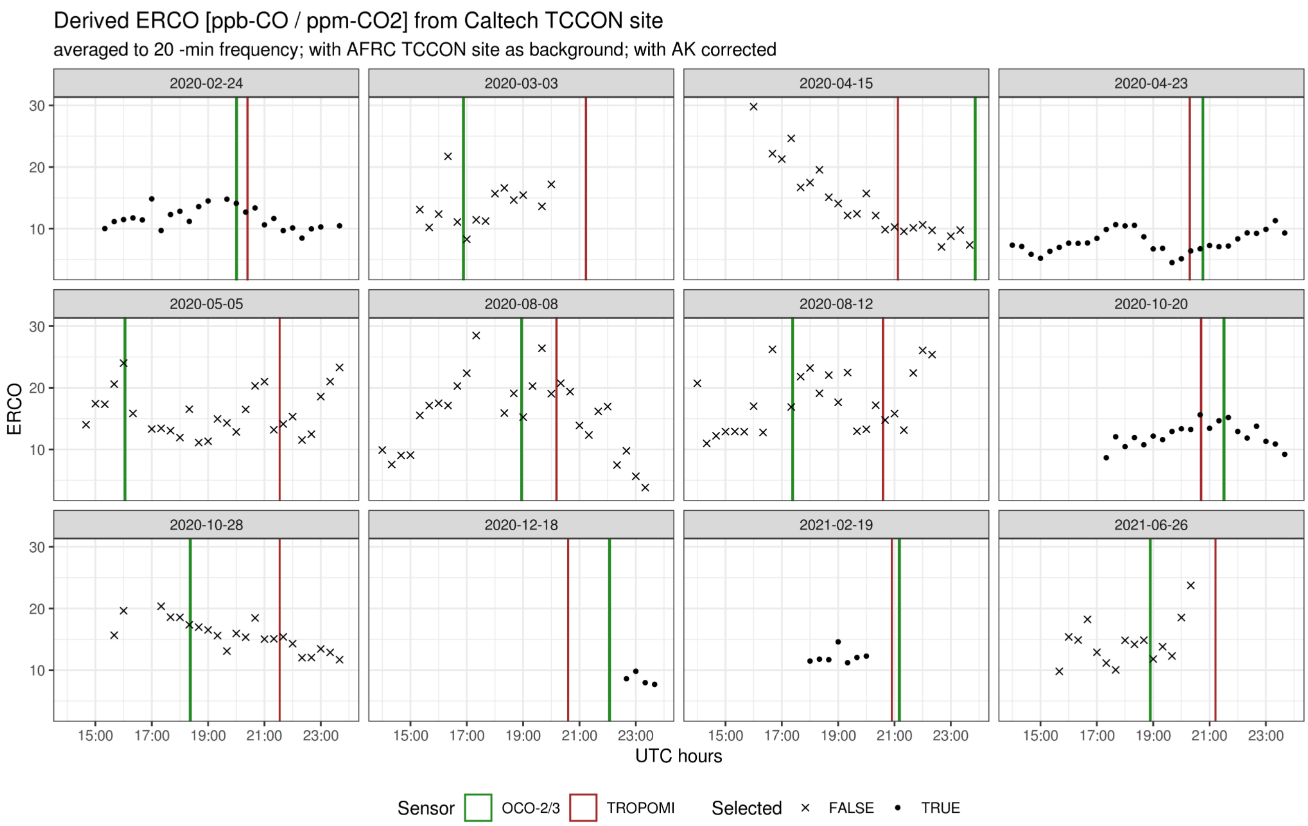


Fig. S8 - Time series of observed ERCO at the California Institute of Technology (Caltech) TCCON site (Wennberg et al. 2017) on the OCO-3 overpass dates till June 2021. The background is defined using the NASA Armstrong Flight Research Center (AFRC) site near Lancaster, California (Iraci et al., 2022). Column enhancements with corrections of TCCON averaging kernel are calculated following Hedelius et al. (2018). The overpasses that went into the final result are shown in solid black dots, while the remaining overpasses with significant plume shift between OCO-3 and TROPOMI overpass times are shown in black crosses. The vertical lines indicate the OCO-3 (green) or TROPOMI (red) overpass times in UTC times. The day of the week for each date is shown as follows: Feb 24, 2020 (Mon), Mar 3 (Tues), Apr 15 (Wed), Apr 23 (Thurs), May 5 (Tues), Aug 8 (Sat), Aug 12 (Wed), Oct 20 (Tues), Oct 28 (Wed), Dec 18 (Fri), Feb 19, 2021 (Fri), and June 26, 2021 (Sat). Note that no qualified data exists during the overpass time of OCO-3 or TROPOMI on April 1, 2020. The TCCON data were obtained from the TCCON Data Archive hosted by CaltechDATA at https://tccondata.org. We thank Laura T. Iraci and Coleen M. Roehl for preparing the TCCON data for these two sites.

Even though many more analyses related to the hourly or seasonal ER_{CO} can be drawn from these TCCON observations, they are out of the main goal of the current manuscript in using space-based observations. Since these TCCON and spatially resolved satellite observations (OCO-3-TROPOMI) may have a different emphasis on the emission signals within the LA basin, we tried to avoid unnecessary direct comparisons of their derived ER_{CO} and decided to leave the analysis in the supplementary material but added some discussions in

Sect. 2.1.3:

3. *Overpass times, meteorological conditions, and emission variations:* As a result of the overpass time difference between sensors, variations in meteorological conditions (e.g., wind direction and speed) can lead to changes in urban plume shapes detected by the two sensors as they pass by. We dealt with changes in wind speed and wind direction separately. The former is resolved by the "scaling factor" inferred from an atmospheric transport model and the latter undergoes manual evaluations (**Sect. 3.1**). Besides, CO and CO_2 emissions themselves can vary over the course of a day, e.g., driven by road transportation and residential sectors. Given the overpass time difference between sensors, it is likely that such mismatch in the timing of CO versus CO_2 emissions may affect the observed ER_{CO}.

Sect. 3.1:

Besides changes in wind directions, CO and CO_2 emissions themselves can vary across daytime hours, likely driven by the road transportation and residential sectors. As a result, variations in the derived ER_{CO} across multiple overpasses may reflect not only the variation in combustion efficiencies but also the mismatch in the emission timing. LA may be one of the cities with more distinct daytime changes in emissions compared to industry-centered cities. Fortunately, based on a supplementary sensitivity analysis using measurements from the Total Carbon Column Observing Network in Pasadena (TCCON, Wennberg et al., 2017), by limiting satellite overpasses to those with a smaller time difference, ER_{CO} appear to be less variable (**Supplementary Fig. S8**). Future geostationary satellite monitoring NO_x (e.g., TEMPO, Chance et al., 2022) may provide better guidance towards the hourly pattern in urban emissions, especially from the traffic sector with more daytime fluctuations, which have been discovered using surface monitoring networks (e.g., over Chicago; de Foy, 2018).

And Sect 4.1:

"If TROPOMI pixel sizes are relatively large (i.e., non-nadir observations) or the wind is steadier, this dt constraint may be relaxed, as long as emissions for a specific city is less driven by sectors with noticeable diurnal cycle (e.g., road transportation)."

Reference:

Hedelius, J. K., Liu, J., Oda, T., Maksyutov, S., Roehl, C. M., Iraci, L. T., Podolske, J. R., Hillyard, P. W., Liang, J., Gurney, K. R., Wunch, D., and Wennberg, P. O.: Southern California megacity CO2, CH4, and CO flux estimates using ground- and space-based remote sensing and a Lagrangian model, Atmos. Chem. Phys., 18, 16271–16291, https://doi.org/10.5194/acp-18-16271-2018, 2018.

Wennberg, P. O., D. Wunch, C. Roehl , J.-F. Blavier, G. C. Toon, N. Allen. 2017. TCCON data from California Institute of Technology, Pasadena, California, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.pasadena01.R0

Iraci, L., J. Podolske, C. Roehl, P. O. Wennberg, J.-F. Blavier, N. Allen, D. Wunch, G. Osterman. 2022. TCCON data from Armstrong Flight Research Center, Edwards, CA, USA, Release GGG2020R0.

TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.edwards01.R0

Minor concerns:

What the authors mean by "FF" needs to be made clear. The implication is that the emission signature of fossil fuels is being measured directly, which is clearly not the case. Emission from combustion, sure. There is no capacity to separate e.g. biofuel from fossil fuels in this approach, and this needs to be made clear.

We agree that FF enhancements need to be more precisely defined. It is challenging to isolate the biofuel signals from fossil fuel signals by only using limited atmospheric observations. Nevertheless, the proportion of biofuel combustion would likely be small compared to fossil fuel combustion. We now added a note for clarifications in the second paragraph of Sect. 2:

"Since we do not differentiate emission signals due to biofuel or fossil fuel (FF) combustion, the term "FF enhancements" is simply referred to *column enhancements induced by any anthropogenic combustion processes* from the target city."

L37: While this is true for some air pollutants such as CO, this is not true for many CO2. In fact, increasing efficiency during combustion activities increases the amount of CO_2 and NO_x emitted (unless the latter is scrubbed) (e.g. Lama et al., 2020). The key to reducing emissions of greenhouse gases is to reduce combustion, period.

We agree with the reviewer that the initial statement was inaccurate. As reviewer #1 and previous studies have also mentioned, NO_x emissions will be higher during efficient combustion with high temperatures. We have now taken this point into account and modified the text as:

"Given the co-benefit between GHG reduction and improved air quality at various scales (e.g., Zhang et al., 2017), controlling the consumption of fossil fuels altogether is the key."

L223-224: Why use a rectangle as the source function? Was this to be "independent of emission inventories", while still knowing that emissions are locating within the city?

A city is simply represented by a rectangle that outlines the broader boundary for possible emissions. The assumption of a rectangle shape is somewhat limited by the model code. Ideally, the spatial extent of urban emissions could be any shape but will then involve more info about the spatial distribution of emissions. To determine a rectangle that covers a wide area, we looked at google map images or urban land cover maps from WUDAPT (if applicable).

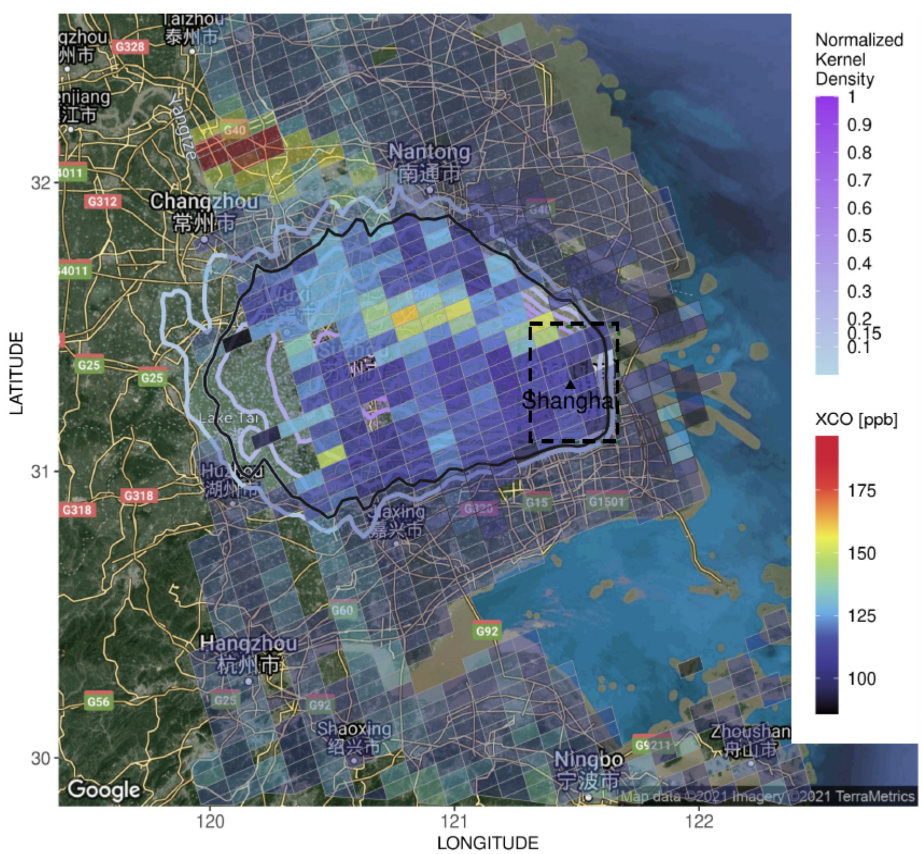
We may argue that it is fair to assume FF emissions are very likely to be located within the city and such an assumption may not contradict the statement of "independent of emission inventories". Emission inventories provide very specific information about the magnitude and spatial distribution of FF emissions. One can think of the other end of the spectrum: we release air parcels from the individual sources according to emission inventories in the cities (e.g., more air parcels are released from the grid with larger emissions) and watch how those air parcels move in space and time. Compared to such an approach, our assumption and approach involve very little info from inventories (no info about specific emission locations nor magnitudes).

Lastly, we would stress that the idea is more about estimating the probability of urban plume realized by the particle distribution during overpass time and the quantitative expression via 2D kernel density.

L231: Regarding the second point about excluding observations elevated by another city: yes, this excludes enhancements outside the model-defined urban plume, *but what about enhancements from other cities that might also be contributing to enhancements within the urban plume?* This seems to be the case for the XCO values shown in Figure 3c. Is there any way to correct for these values? Especially given that there do not seem to be XCO2 values over the same area?

Enhancements within the urban plume might be influenced by other cities. It is possible to calculate the FF enhancements due to other cities' emissions, for example, using footprint and prior emissions. Or, one can perform the full atmospheric inversion to optimize emissions from not only Shanghai but also nearby cities. But, such calculations or inversion will involve prior assumptions, which is less wanted for this work. Alternatively, we could further limit the spatial extent of the urban plume by choosing a higher kernel density. However, we may argue that the current setting is sufficient for the cases we examined. As the reviewer mentioned, there seem to be no XCO_2 values in those regions. Since ER_{CO} was only calculated when OCO-3 XCO_2 is available, there seems to be a lower priority for such corrections.

c) TROPOMI, Feb 20th, 2020, 04:44 UTC



L235: Is there any way to mark on here which soundings were used to define the background? Some more quantification of the information would also be useful here. Which latitude range was used to define the background for the OCO-2 swath? In Figure 3a I don't see any soundings outside the urban plume region… For 3b it is somewhat clearer, aided by the coastal cut-off (although point sources near the coast are clearly in the “background”, but how the area would be defined for Figure 3c is really vague. With this level of detail, the approach would be hard to replicate.

The background soundings/polygons were actually highlighted with a black outline (e.g., polygons to the east of the urban plume), while other soundings/polygons were outlined in white. However, the color seemed to be difficult to observe as the reviewer pointed out. We have now added several notes in the figure caption to guide readers.

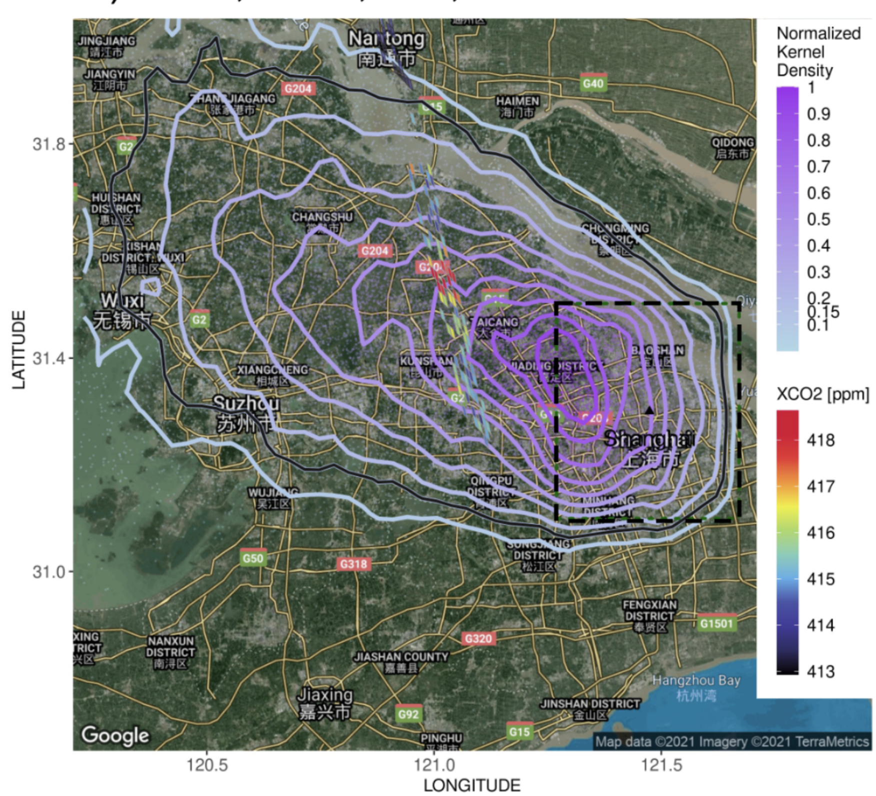
For clarification, we allowed for a 0.1degree “buffer area” outside the urban plume (to further avoid including some enhanced values, even though a random wind error has been included to generate the urban plume). The background range/domain is further extended up to 0.3 degrees in either longitudinal or latitudinal direction depending on the wind direction. We also tested a background range of 0.5 degrees, but the resultant background values did not change much. The background extension range is a free parameter in the X-STILT background r code, where the user can test their own values for their targets.

Here are more details and explanations regarding the reviewer’s individual questions.

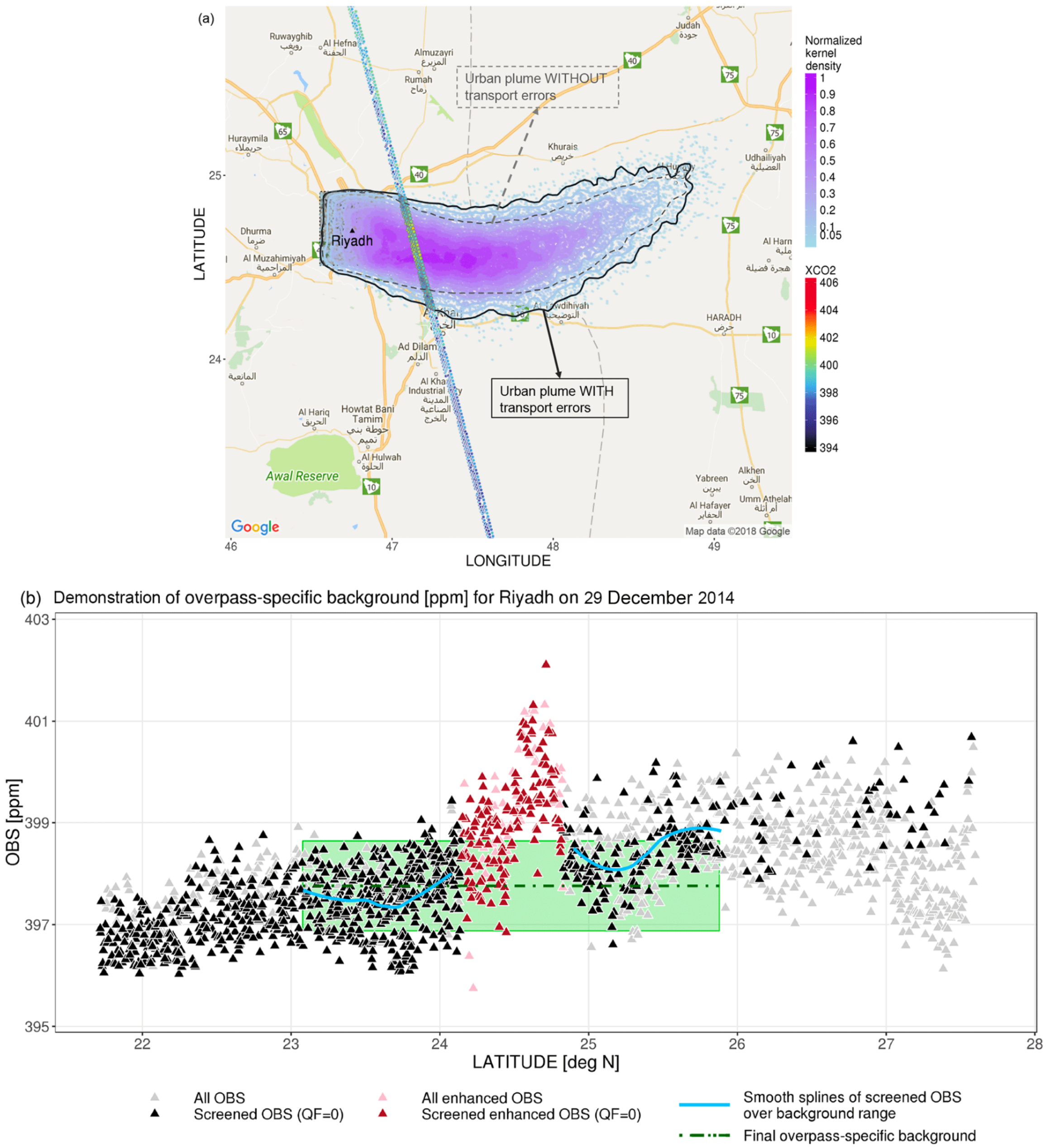
- For Fig. 3a, there were a couple of soundings to the north outside the urban plume (around 121.0 degE, 31.9 degN). Because those background soundings are tilted with relatively lower values (displayed in dark blue), they could be hard to identify. We have now added a note in the figure caption to guide readers.

  Ideally, we would choose OCO-2 soundings to the east, or the south of the urban plume given the southeasterly wind on Feb 4th, 2020, but the narrow swath of OCO-2 makes it impossible. Hence, we had to treat either the northern or southern soundings as background soundings.

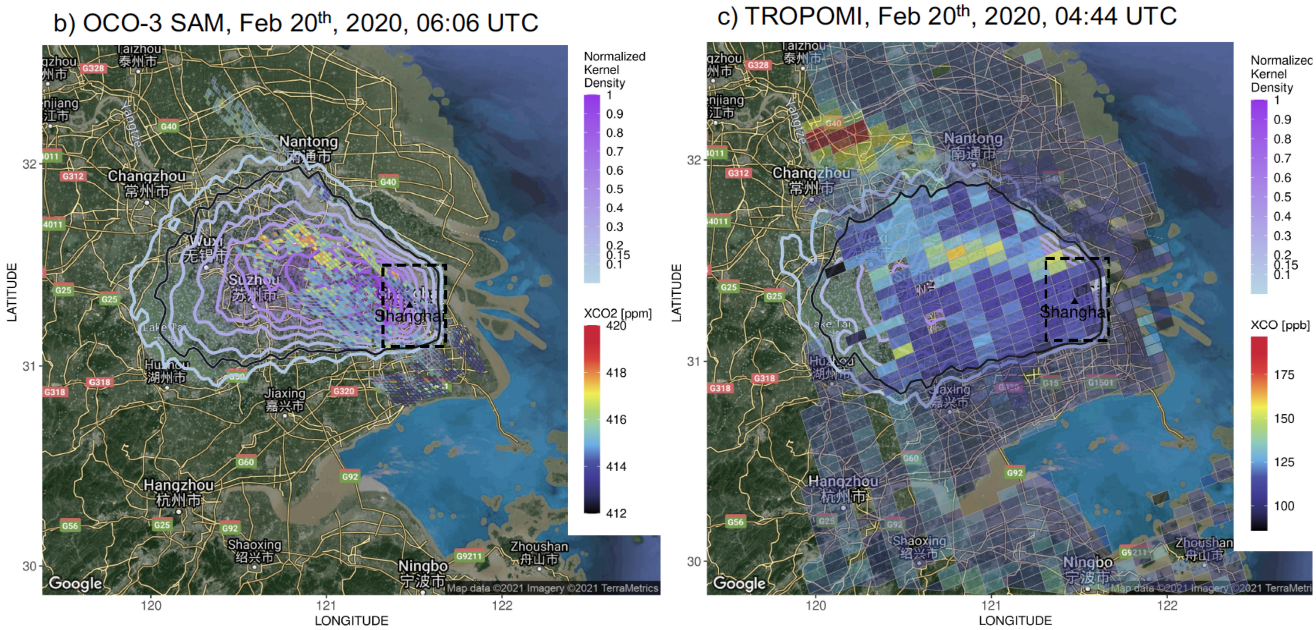
a) OCO-2, Feb 4th, 2020, 05:11 UTC



As explained in Wu et al. (2018), we showed a better demonstration of the background using OCO-2 data over Riyadh (https://gmd.copernicus.org/articles/11/4843/2018/#&gid=1&pid=1, see figure copied below). A wider latitude range of 1 degree was chosen in that Riyadh case because Riyadh is a relatively isolated city as we discussed in that paper. While for Shanghai (surrounded by other Chinese cities), we limited the background range to 0.3 degrees to avoid potential influence from a nearby city and tested an alternative of 0.5 degrees for the impact.



- For Fig. 3b (in the figure below), the background soundings were also highlighted in a darker color to the south of the urban plume (latitude < 31 degN). Individual swaths have their own slightly different background values. One may also see there is a small gap between the urban plume in the solid black curve and the darker soundings with latitudes < 31 degN, corresponding to the “buffer area” mentioned earlier.



- For Fig. 3c, the background soundings are highlighted with a black outline rather than a white outline for other soundings outside the urban plume. And, the soundings within the urban plume are outlined in grey. As explained above, the longitude range of the background region extends from around 121.7 degE toward the ocean (~ 0.3 degrees along longitude). The latitude range of the background region has the same range as that of the urban plume (i.e., from ~30.8degN to 31.9degN). These explanations have been added to the figure caption to guide readers.

Section 2.2.4: It's unclear how the observation uncertainty is computed here. Each sounding has a reported uncertainty that comes with the data product – is this the retrieval error? The "measurement noise" seems to be equated with the standard deviation of the retrievals within a TROPOMI footprint. But this isn't really what is usually meant by measurement noise.

We apologize for the lack of details on uncertainty calculations. Yes - each sounding is associated with a reported retrieval posterior XCO_2 uncertainty from Level 2 files. We agree with the reviewer that the terminology of "measurement noise" is not that accurate. As investigated by Worden et al. (2017), the standard deviation (SD) of retrieved XCO_2 values contains contributions from several possible sources, including 1) natural variability, 2) measurement noise and 3) a slowly varying bias (Sect. 3 of https://authors.library.caltech.edu/85361/1/amt-10-2759-2017.pdf).

The SD of XCO_2 retrieval is still needed and better be termed as "uncertainties due to binning", since we had to reconcile the resolution difference between two satellites, by taking the average of observed $XCO_{2.ff}$ enhancements from a few OCO-2/3 soundings within a given TROPOMI polygon to arrive at the averaged $XCO_{2.ff}$ enhancement at the TROPOMI scale. This "binning or average uncertainty source" was also reported in Table 1 by Wu et al. (2018).

It is not completely clear how these two types of errors (one per OCO-2 sounding and one per larger TROPOMI sampling) are combined, only that the retrieval errors are aggregated in a "standard- deviation-of-mean" manner, which sounds like it's being divided by the square root of the number of soundings. *This assumes that the measurements and their errors are independent of one another, which is not the case.*

This almost certainly underestimates the measurement uncertainty on the X_{ffCO2} term. A formula here is certainly needed.

We agreed that the lack of error covariance between every OCO-3 sounding within a TROPOMI grid may underestimate the observational uncertainty for the FFCO2 enhancement part. We have now performed additional variogram analyses to obtain a typical error correlation length scale (Lx, km) for all the overpasses per city. There were spatial gaps between OCO-3 soundings for some overpasses, leading to a higher correlation length scale. Thus, we adopted the minimum Lx among all overpasses as the final Lx per city (~normally around 7 to 15 km when using 4 to 6 bins). Below is an example of the variogram analysis to *investigate how retrieval errors correlate in space*.

Interestingly, including the extra covariance term of the retrieval errors leads to a small impact on the total observational error variance, because the overall observational error variance is often dominated by the background error component.

To address the reviewer's comments on error analysis, we have rewritten Sect. 2.2.4 and added Fig S4 for the variogram analysis:

2.2.4 Uncertainty sources

The uncertainty related to emissions should contain uncertainties from 1) the atmospheric transport (i.e., column footprints), 2) observations, and 3) non-FF sources and sinks, according to **Eqs. 1 or 2**. We neglect uncertainties from column footprints assuming no transport bias exists during either OCO or TROPOMI overpass time. The urban-background gradient in non-FF fluxes remains very small compared to FF enhancements (**Sect. 3.1**).

We estimated uncertainties of observed FF enhancements following **Eq. 5**. As previously described, observations from a few screened OCO soundings ($\sim$ 5 to 28 OCO soundings depending on the TROPOMI footprint size) are averaged to arrive at mean XCO_2 at the TROPOMI scale. Due to such averaging/binning process, the XCO_2 uncertainty due to binning is considered using the standard deviation of XCO_2 observations ($\sigma^2_{\varepsilon,bin}$ in **Eq. 5**) within a TROPOMI polygon.

$$\sigma^2_{\varepsilon,obs} = \sigma^2_{\varepsilon,bin} + \sigma^2_{\varepsilon,bg} + \sigma^2_{\varepsilon,retrv}. \qquad (5)$$

$\sigma^2_{\varepsilon,bin}$ is not required for estimating XCO uncertainty. Background uncertainty ($\sigma^2_{\varepsilon,bg}$) contains both the retrieval error and the variability of column observations (as standard deviation) within background regions.

The retrieval uncertainty ($\sigma^2_{\varepsilon,retrv}$) of XCO is available for each TROPOMI sounding, whereas that of XCO_2 is reported for individual OCO-2/3 sounding (as read from Level 2 Lite files), which need to be aggregated at the TROPOMI scale. Due to possible correlations in retrieval errors between nearby OCO soundings, we estimated the error correlation length scale (L_x) using exponential variograms as demonstrated in **Supplementary Fig. S4**. Within a TROPOMI polygon that contains N numbers of OCO soundings, an error variance-covariance matrix with a dimension of $N \times N$ is constructed with its diagonal elements filled with OCO sounding-specific retrieval error variance. Then, L_x is used to form the normalized covariance matrix, i.e., $exp(-\frac{D(S_i,S_j)}{L_x})$ where $D(S_i, S_j)$ denotes the distance between each two OCO soundings ($1 \leq i < j \leq N$). Lastly,

the sum of all elements in the error covariance matrix (both variance and covariance elements) is divided by N^2 to obtain one $\sigma^2_{\varepsilon,retrv}$ per TROPOMI grid. As a result, the overall uncertainty of FF enhancement per sounding is often predominated by the background error component.

Here is the new Fig. S4 for an example of the variogram analysis to examine horizontal correlation of the retrieval XCO_2 errors (not retrieved XCO_2 values for clarification):

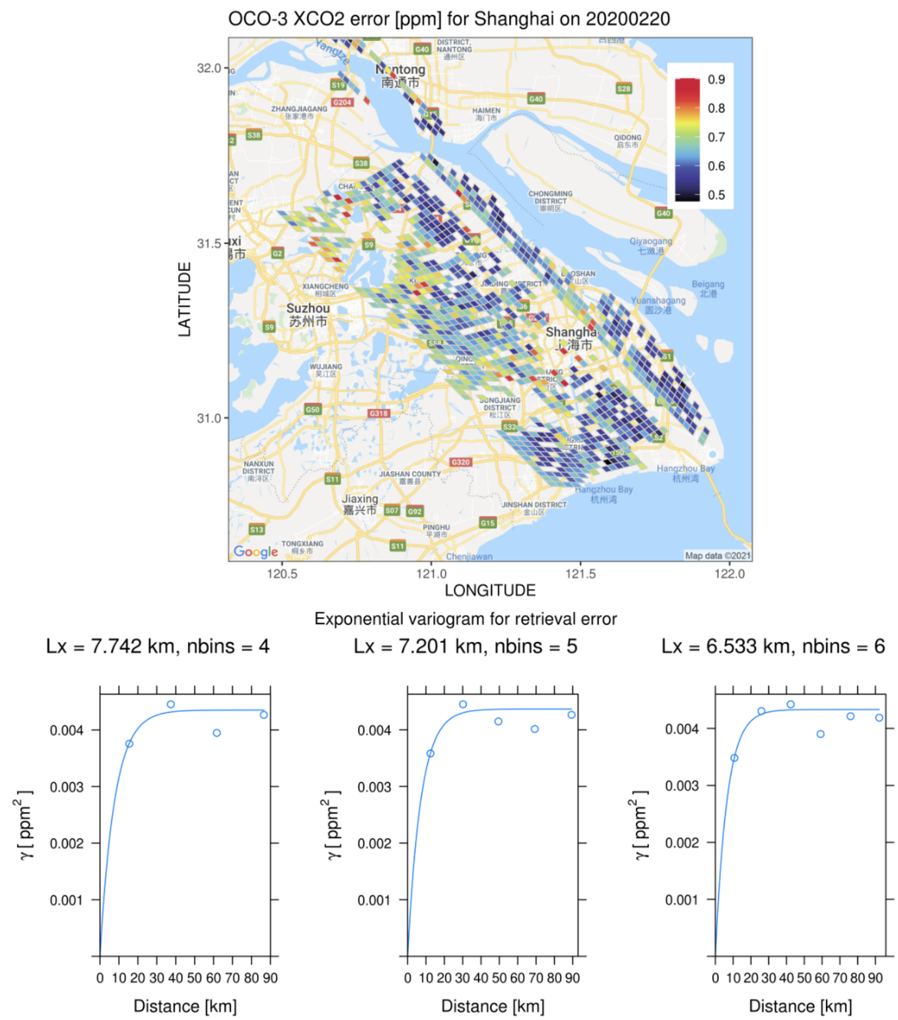


Figure S4. **Upper Panel**: An example of the sounding-specific XCO_2 retrieval errors from L2 files over Shanghai on Feb 20, 2020. **Lower Panels**: Exponential variogram analyses of retrieval errors and resultant error correlation length scale (L_x, km) with 4, 5, or 6 bins.

Technical/language comments:

L43-45: I'm confused by this sentence. Maybe should "Benefit" be "Benefiting"? In any case, the sentence should be rewritten to make it clearer, or even split into two sentences, starting with the second half, i.e. something like "The ratio… between tracers is reported. This has the benefit that errors … cancel out."

The text has been reworded as -

"The commonly-used approach in estimating combustion efficiency is to combine atmospheric observations of multiple trace gases and to report the ratio of the total or excess measured concentrations (above a defined background value) between tracers \citep{Silva2017,Reuter2019,Park2021ratio}. Such tracer-to-tracer ratio calculation has the benefit that errors in describing the atmospheric transport that carries tracers to the measurement site can be canceled."

L46: would remove “their”

L50: …difficult to detect?

L53-54: Rewrite this, something like: “Given its much longer lifetime, CO is much easier to interpret…” The “on the other hand” doesn’t fit here as written.

Figure 1 caption: “The x-axis indicates…”

L82: such gradient -> either “such a gradient” or “this gradient”

L87-88: in vertical -> in the vertical

L89: by the gaps in prior literature -> by gaps in the existing literature L92: adopt or adapt?

L97: quantify accurately -> accurately quantify

L101: implication and limitation -> implications and limitations

L107: surrounded -> surrounding

L109: urban plume that is the spatial extent -> urban plume, defined as the area

L111: that estimated -> that are estimated

L113: requires estimate -> requires an estimate

L116: would remove “much”

L155: accounted for AK -> accounted for the AKs?

L156: As result -> As a result

L160: evaluations -> evaluation

L163: interfere -> interfere with; also: more explanations in Sect. -> for more explanation, see Sect.

> We appreciate the careful read-through and detailed comments from the reviewer. Grammatical issues associated with **L46 to L163** have been corrected as suggested.

Figure 2 caption: The citation for the Google Maps data in the last sentence doesn’t sound quite right – please check what it’s supposed to be (i.e. adopted the Google Maps what?)

> Revised as: “The underlying hybrid maps were created using the ggmap library in R with the hybrid view of Google Maps over LA (copyright: Map data \copyright 2021 Imagery \copyright 2021 TerraMetrics).”

L166: and atmospheric transport model -> and an atmospheric transport model

L175: for sounding-specific -> for the sounding-specific

> Both corrected.

L178-181: I think I understand what is meant here, but it’s a bit hard to parse. When I hear “pathways” I’m thinking of chemical reactions, and I’m not sure what is meant by an air parcel being “tied” to somewhere, or correction terms being “attached” (perhaps “applied” would fit better)? In any case this should be rewritten for clarity.

> We have changed “pathways” to “routes”. Here is the revised sentence: “From a Lagrangian viewpoint, the air parcels arriving at an urban sounding might be traced back to different origins from the air parcels arriving at a rural sounding, meaning observations at the two soundings may be influenced differently by the surrounding biosphere.”

L193: corresponding for -> corresponding to

L196: wind condition -> wind conditions

L200: by sounding -> by the sounding

L201: Because AK -> Because the AK
L204: If -> By

All corrected.

L208-209: I guess the meteorological conditions and the AK profile are specific to the sounding, not just the sensor?

Correct - we have changed to “sounding-specific”.

L209: condition -> conditions; L216: combines-> combine

Corrected.

L216-217: This seems a bit backwards – isn’t the first method more independent of information about emissions, unlike the two modelling-based approaches?

Correct - this statement has now been removed. For clarification,

- The first statistical method utilized standard deviation or percentiles to derive the background purely from observed XCO_2. So, no use of prior emissions.
- The second method uses modeled initial conditions from atmospheric transport models and global CO_2 fields. Global CO_2 fields may involve prior assumptions on emissions.
- The third method first considered atmospheric transport and identified the background region (but not too far from the target city) and then calculated the median of the observed XCO_2 in the background region as the background. An assumption on the rough spatial extent of the city emission (e.g., a rectangle). Please refer to our response above regarding this assumption.

L218-219: improve it over what?

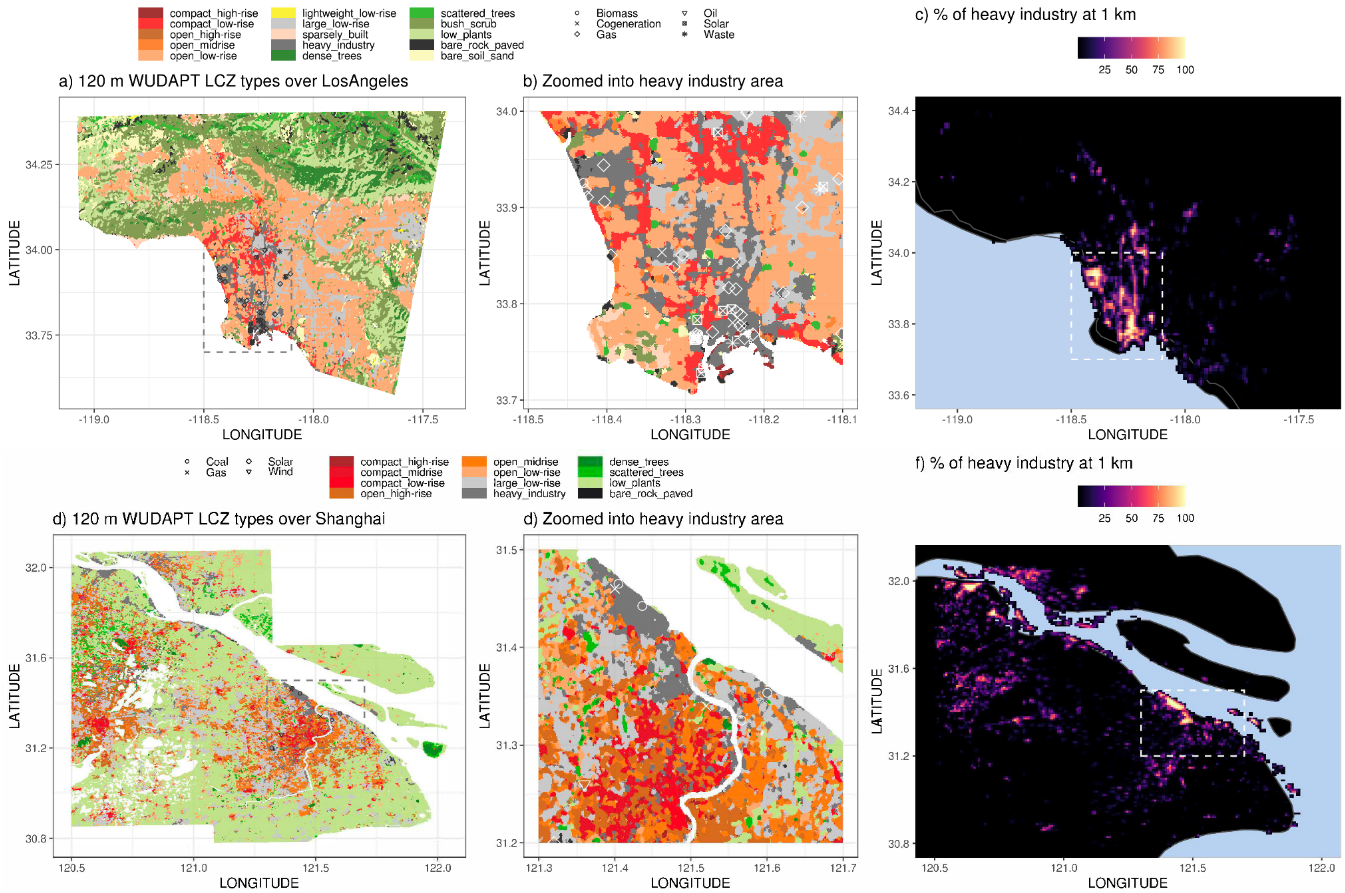
Worded as - “The broader spatial coverage and multiple swaths stretching out of the city domain of OCO-3 SAMs improve the determination of the background terms by introducing spatial variations in the background, compared to the narrow swath of OCO-2.”

L222: soundings within -> soundings as within
L258-262: I would split this into two sentences.
L265: lasts for -> ranges from; L267: are -> is
L273: Observed uncertainty of XCO2 are -> Observation uncertainty of XCO2 is

All corrected.

Figure 4: Please add a coastal outline in panels c and f to make it easier to interpret. (Masking water would also be an option.); Figure 4 caption: lightgray should be two words.

Corrected. Here is the updated Figure 4 with the ocean colored in light blue:



L301: I think a word is missing. Maybe: “Those industry coverage *maps* are then convolved…”

Corrected.

L311: I would remove “originating” here, it is more confusing than helpful.

Corrected.

L313: When you write “too low valid soundings” do you mean “too few valid soundings” or “too low enhancements”?

Corrected - too few valid soundings.

L313-314: remove “the few”

L319: interfered by wind shift -> affected by shifting winds

L329-330: Recommended change: “Again, the colored contours and curves in Fig. 5 indicate neither the intensity of concentrations nor flux fields (as no prior emissions are used), but rather the likelihood of urban plumes determined solely by atmospheric dispersion.”

All corrected. It reads much better.

L332: but problematic -> but becomes problematic

L335: remove “cases”

Figure 5: in figure label, it should be “caution” rather than “cautions”

Corrected.

L364: I would suggest replacing “atmospheric movement” with “transport”. Also, something seems to have gone wrong with many of the subscripts in this paragraph (e.g. ECO instead of E$_{CO}$ in LaTex syntax).

Subscripts of E_{CO} and E_{CO_2} have been corrected.

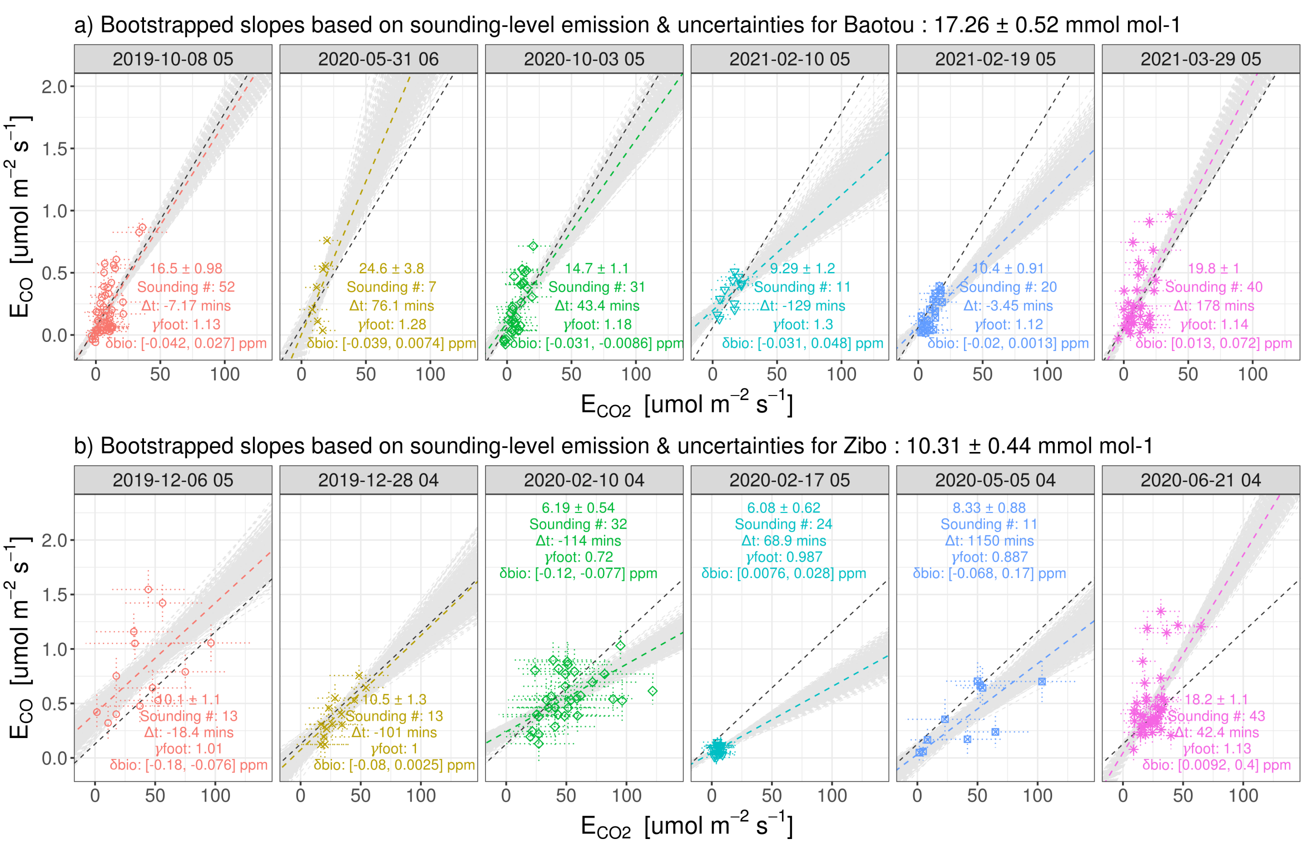
Figure 6 caption: The second sentence and the second-last sentence seem to repeat the same information. Remove one? Also, it is mentioned here that only regression lines with positive slopes were chosen from the Monte Carlo experiment. What proportion of these lines needed to be removed?

Caption corrected. Yes - the Monte Carlo analysis will remove bootstrapped linear fits with negative slopes, assuming CO and CO_2 emissions are positively correlated at the city scale (when sufficient soundings are being considered). Since it is a random process, the exact proportion of lines with negative slopes may vary each time. We added a note in the figure caption of Fig. 6:

> “Only bootstrapped regression lines with positive slopes were chosen from the Monte Carlo experiment (dashed gray lines; ~98.4% and 93.3% of the total 6,000 bootstrapped lines for Baotou and Zibo, respectively).”

Figure 6: the \gamma bio term shown on the plots should be in units of ppm, right? Also, is this somehow different from the \delta X_{bio} discussed previously? If not, please make this consistent.

We modified the two figures and text to remove \gamma bio, to reduce the number of terminologies introduced in the paper. Instead, we now printed the range of \delta X_{bio} in the interval format (i.e., min and max values of the urban-rural gradient in biogenic anomalies estimated at the TROPOMI scale, in ppm). Here is an example of Zibo and Baotou cases (please see the last row of text in each panel):



L385: remove second comma, also change “them” to “these overpasses”.
L392: in time -> in the time
L406: tend -> tends
L408 (and elsewhere): perhaps “industry-dominated” might be more appropriate in some places than “industry-dominant”?
L432: i.e., less -> i.e., those less
L452: the Solar-Induced -> remove “the”
L455: function -> functional; L456: either “interfering” -> “interfering with” or “interfering” -> “affecting”; “wind directional shift induced by” -> “the shift in wind direction due to”
L464: spontaneously -> simultaneously
L464: do you mean “a future geostationary satellite (i.e. GeoCarb)” or “future geostationary satellites”?

All corrected. We meant “future geostationary satellites”.

Figure 8 caption: Text could use a bit of work. Suggestion: “indicate the urban plumes between two times differ significantly that a simple plume rotation fails to fix” -> “indicate that that urban plumes between the two overpass times differ so much that they cannot be brought into agreement with a simple plume rotation”. Also: what is the meaning of 0~1 and 0~2? Are they shifted or not?

Corrected. We added a note in the caption to explain “0~1”: “The number on each bar (e.g., 0~1) denotes the number of TROPOMI polygons needed to be shifted to align the urban plumes at two times. For example, 0~1 means that TROPOMI polygons over certain locations are shifted by one grid.”

L479: begun -> began; L483: take -> takes
L492: to coarser -> to the coarser
L494: may locate -> may be located
L500: against OH -> against the OH
L503: contributed to 21.2% -> contributed 21.2%
L503: but negligible -> but a negligible
L504: season, the -> season, and as such the
L504: remove “likely”, “encapsulated” -> “included”, “yield” -> “have”
L506-507: “…whether … remains unclear” -> “it is unclear whether AVOCs… of interest.”
L512: “can help” -> “to”
L513-514, and L521: inventory -> inventories
L515: footprint -> footprints
L520:
L527: in informing locations of -> to provide information about

All corrected.

Supplement:

Figure S1 caption: average -> averaged

Corrected.

Figure S3 caption: unique to each satellite sounding given unique -> unique to each satellite sounding, giving a unique; Figure S3 caption: Column footprint -> The column footprint

Corrected.

Figure S4 caption: “these resultant normalized fraction Pind(x, y) informs the influence on the observation at a given sounding (white rectangle) due to heavy industry. Lighter the color, stronger impact from heavy industry in LA.” -> “these resultant normalized fractions Pind(x, y) show the influence of heavy industry on a given sounding (white rectangle). The lighter the color, the stronger impact from heavy industry.”

Corrected.

Figure S5 caption: “of urban-background” -> “of the urban-background”; “Since biogenic” -> “Since the biogenic”; “Two sets of footprint totals” -> “The two sets of footprints”; “difference in AK” –> “difference in the AKs”; “Higher the footprint ratio, larger the discrepancies” -> “The higher the footprint ratio, the larger the discrepancy”

Corrected.

Figure S6 caption: “during two” -> “during the two”

Corrected.

Figure S7 (and others): Is the date format in the plots (YYYYMMDD) consistent with the Copernicus guidelines? Figure S7 caption: “close to the noon, daytime carbon sink dominant leading” -> “close to noon, the daytime carbon sink dominates, leading”; “nighttime respiration dominant,” -> “respiration dominates the biogenic fluxes,” (I would remove “nighttime” because winter cases are also mentioned.)

Fig. S7 has been revised to show the dependence of overpass time/hour, according to a minor comment from reviewer #1. And the date format has been modified -

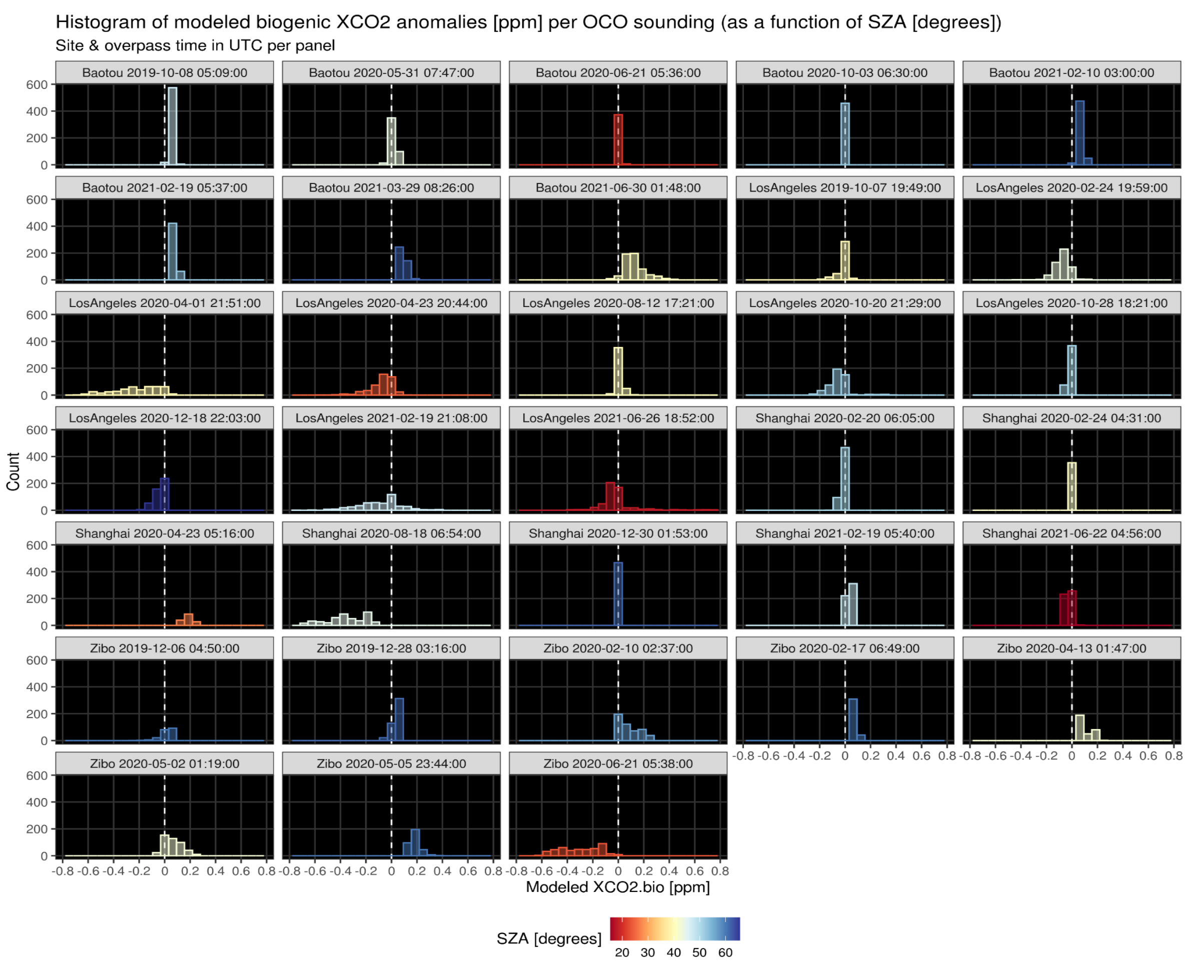


Figure S9 caption: "with or without the account of the urban-rural biogenic gradient over Zibo on June 21, 2020" -> "over Zibo on June 21, 2020, with and without taking the urban-rural biogenic gradient into account"; "light grey shading denote" -> "light grey shading denotes"; "Such positive gradient" -> "Such positive gradients"

Corrected.

Figure S10 title: A bit awkward, would suggest "Log-normal distributions of ERCO [ppb-CO / ppm/CO2]; Figure S10 caption: "on log-normal" -> "on a log-normal"; "stewed" -> "skewed"; Also, aren't the values skewed towards the higher end (i.e. positive skew or right-skewed) if the mean is higher than the median?

Right - right-skewed when mean is larger than median. The title and caption have been corrected.

Figure S11: Should the title for panels c) through f) be Sectoral CR_CO or ER_CO? Also, I guess panel f) should be marked "residential" rather than "resident"? (Perhaps "road traffic" might also be better for d)…)

Right - those panels c) through f) are simulated enhancement ratios using EDGAR-based sectoral emissions. Just for clarification, here are the FF sub-sectors from EDGAR we selected and how we modified them.

 # ENE: Power industry ***-> renamed as 'power' (Fig. S11e)***
 # TRO: Road transportation ***-> renamed as 'on-road' (Fig. S11d)***
 # RCO: Energy for buildings ***-> renamed as 'resident' (Fig. S11f)***

Subsectors that were ***combined into one sector as "industry" (Fig. S11c)***
 # IND: Combustion for manufacturing
 # CHE: Chemical processes
 # NMM: Non-metallic minerals production
 # NFE: Non-ferrous metals production;
 # IRO: Iron and steel production

We have now modified panel titles for the bottom panels.